

1644

RAW SEQUENCE LISTING
PATENT APPLICATION: US/09/403,882A

DATE: 05/10/2001
TIME: 11:41:49

Input Set : A:\ES.txt
Output Set: N:\CRF3\05102001\I403882A.raw

ENTERED

3 <110> APPLICANT: University of California, San Francisco
4 Farinas, Javier
6 <120> TITLE OF INVENTION: Methods and Reagents for Targeting Organic Compounds To Selected Cellular Locations
7
9 <130> FILE REFERENCE: UCSF1100-3
11 <140> CURRENT APPLICATION NUMBER: 09/403,882A
12 <141> CURRENT FILING DATE: 2000-03-20
14 <150> PRIOR APPLICATION NUMBER: 60/081,340
15 <151> PRIOR FILING DATE: 1998-04-09
17 <150> PRIOR APPLICATION NUMBER: 60/081,118
18 <151> PRIOR FILING DATE: 1998-04-08
20 <160> NUMBER OF SEQ ID NOS: 9
22 <170> SOFTWARE: PatentIn version 3.0
24 <210> SEQ ID NO: 1
25 <211> LENGTH: 951
26 <212> TYPE: DNA
27 <213> ORGANISM: Artificial sequence
29 <220> FEATURE:
30 <223> OTHER INFORMATION: single chain antibody
32 <220> FEATURE:
33 <221> NAME/KEY: CDS
34 <222> LOCATION: (1)..(951)
36 <400> SEQUENCE: 1
37 atg gcc gag gtc aag ctg cag gag tca ggg gga ggc tta gtg cag cct 48
38 Met Ala Glu Val Lys Leu Gln Glu Ser Gly Gly Gly Leu Val Gln Pro
39 1 5 10 15
41 gga ggg tcc cgg aaa ctc tcc tgt gca gcc tct gga ttc act ttc agt 96
42 Gly Gly Ser Arg Lys Leu Ser Cys Ala Ala Ser Gly Phe Thr Phe Ser
43 20 25 30
45 agc ttt gga atg cac tgg gtt cgt cag gct cca gag aag ggg ctg gag 144
46 Ser Phe Gly Met His Trp Val Arg Gln Ala Pro Glu Lys Gly Leu Glu
47 35 40 45
49 tgg gtc gca tat att agt agt ggc agt agt acc atc tac tat gca gac 192
50 Trp Val Ala Tyr Ile Ser Ser Gly Ser Ser Thr Ile Tyr Tyr Ala Asp
51 50 55 60
53 aca gtg aag gga cga ttc acc atc tcc aga gac aat ccc aag aac acc 240
54 Thr Val Lys Gly Arg Phe Thr Ile Ser Arg Asp Asn Pro Lys Asn Thr
55 65 70 75 80
57 ctg ttc ctg caa atg acc agt cta agg tct gag gac acg gtc atg tat 288
58 Leu Phe Leu Gln Met Thr Ser Leu Arg Ser Glu Asp Thr Val Met Tyr
59 85 90 95
61 tac tgt gca aga gat tac ggg gct tat tgg ggc caa ggg acc acg gtc 336
62 Tyr Cys Ala Arg Asp Tyr Gly Ala Tyr Trp Gly Gln Gly Thr Thr Val
63 100 105 110
65 acc gtc tcc tca ggt gga ggc ggc tca ggc gga ggt ggc tct ggc ggt 384
66 Thr Val Ser Ser Gly Gly Gly Ser Gly Gly Gly Ser Gly Gly
67 115 120 125

RAW SEQUENCE LISTING

DATE: 05/10/2001

PATENT APPLICATION: US/09/403,882A

TIME: 11:41:49

Input Set : A:\ES.txt

Output Set: N:\CRF3\05102001\I403882A.raw

```

69 ggc gga tcg gac att gag ctc acc cag tct cca gca atc atg tct gca      432
70 Gly Gly Ser Asp Ile Glu Leu Thr Gln Ser Pro Ala Ile Met Ser Ala
71      130      135      140
73 tct cca ggg gag agg gtc acc atg acc tgc agt gcc agt tca agt gta      480
74 Ser Pro Gly Glu Arg Val Thr Met Thr Cys Ser Ala Ser Ser Ser Val
75 145      150      155      160
77 agg tac atg aac tgg ttc caa cag aag tca ggc acc tcc ccc aaa aga      528
78 Arg Tyr Met Asn Trp Phe Gln Gln Lys Ser Gly Thr Ser Pro Lys Arg
79      165      170      175
81 tgg att tat gac aca tcc aaa ctg tct tct gga gtc cct gct cgc ttc      576
82 Trp Ile Tyr Asp Thr Ser Lys Leu Ser Ser Gly Val Pro Ala Arg Phe
83      180      185      190
85 agt ggc agt ggg tct ggg acc tct tac tct ctc aca atc agc agc atg      624
86 Ser Gly Ser Gly Ser Gly Thr Ser Tyr Ser Leu Thr Ile Ser Ser Met
87      195      200      205
89 gag gct gaa gat gct gcc act tac tac tgc cag cag tgg agt agt aac      672
90 Glu Ala Glu Asp Ala Ala Thr Tyr Tyr Cys Gln Gln Trp Ser Ser Asn
91      210      215      220
93 cca ctc acg ttc ggt gct ggg acc aag ctg gag ctg aaa cgg gcg gcc      720
94 Pro Leu Thr Phe Gly Ala Gly Thr Lys Leu Glu Leu Lys Arg Ala Ala
95 225      230      235      240
97 gca gaa caa aaa ctc atc tca gaa gag gat ctg aat ggg gcc gtc gac      768
98 Ala Glu Gln Lys Leu Ile Ser Glu Glu Asp Leu Asn Gly Ala Val Asp
99      245      250      255
101 gaa caa aaa ctc atc tca gaa gag gat ctg aat gct gtg ggc cag gac      816
102 Glu Gln Lys Leu Ile Ser Glu Glu Asp Leu Asn Ala Val Gly Gln Asp
103      260      265      270
105 acg cag gag gtc atc gtg gtg cca cac tcc ttg ccc ttt aag gtg gtg      864
106 Thr Gln Glu Val Ile Val Val Pro His Ser Leu Pro Phe Lys Val Val
107      275      280      285
109 gtg atc tca gcc atc ctg gcc ctg gtg gtg ctc acc atc atc tcc ctt      912
110 Val Ile Ser Ala Ile Leu Ala Leu Val Val Leu Thr Ile Ile Ser Leu
111      290      295      300
113 atc atc ctc atc atg ctt tgg cag aag aag cca cgt tag      951
114 Ile Ile Leu Ile Met Leu Trp Gln Lys Lys Pro Arg
115 305      310      315
118 <210> SEQ ID NO: 2
119 <211> LENGTH: 316
120 <212> TYPE: PRT
121 <213> ORGANISM: Artificial sequence
122 <220> FEATURE:
123 <223> OTHER INFORMATION: single chain antibody
124 <400> SEQUENCE: 2
128 Met Ala Glu Val Lys Leu Gln Glu Ser Gly Gly Gly Leu Val Gln Pro
129 1      5      10      15
132 Gly Gly Ser Arg Lys Leu Ser Cys Ala Ala Ser Gly Phe Thr Phe Ser
133      20      25      30
136 Ser Phe Gly Met His Trp Val Arg Gln Ala Pro Glu Lys Gly Leu Glu
137      35      40      45

```

RAW SEQUENCE LISTING DATE: 05/10/2001
 PATENT APPLICATION: US/09/403,882A TIME: 11:41:49

Input Set : A:\ES.txt
 Output Set: N:\CRF3\05102001\I403882A.raw

```

140 Trp Val Ala Tyr Ile Ser Ser Gly Ser Ser Thr Ile Tyr Tyr Ala Asp
141      50      55      60
144 Thr Val Lys Gly Arg Phe Thr Ile Ser Arg Asp Asn Pro Lys Asn Thr
145 65      70      75      80
148 Leu Phe Leu Gln Met Thr Ser Leu Arg Ser Glu Asp Thr Val Met Tyr
149      85      90      95
152 Tyr Cys Ala Arg Asp Tyr Gly Ala Tyr Trp Gly Gln Gly Thr Thr Val
153      100     105     110
156 Thr Val Ser Ser Gly Gly Gly Gly Ser Gly Gly Gly Gly Ser Gly Gly
157      115     120     125
160 Gly Gly Ser Asp Ile Glu Leu Thr Gln Ser Pro Ala Ile Met Ser Ala
161      130     135     140
164 Ser Pro Gly Glu Arg Val Thr Met Thr Cys Ser Ala Ser Ser Ser Val
165 145     150     155     160
168 Arg Tyr Met Asn Trp Phe Gln Gln Lys Ser Gly Thr Ser Pro Lys Arg
169      165     170     175
172 Trp Ile Tyr Asp Thr Ser Lys Leu Ser Ser Gly Val Pro Ala Arg Phe
173      180     185     190
176 Ser Gly Ser Gly Ser Gly Thr Ser Tyr Ser Leu Thr Ile Ser Ser Met
177      195     200     205
180 Glu Ala Glu Asp Ala Ala Thr Tyr Tyr Cys Gln Gln Trp Ser Ser Asn
181      210     215     220
184 Pro Leu Thr Phe Gly Ala Gly Thr Lys Leu Glu Leu Lys Arg Ala Ala
185 225     230     235     240
188 Ala Glu Gln Lys Leu Ile Ser Glu Glu Asp Leu Asn Gly Ala Val Asp
189      245     250     255
192 Glu Gln Lys Leu Ile Ser Glu Glu Asp Leu Asn Ala Val Gly Gln Asp
193      260     265     270
196 Thr Gln Glu Val Ile Val Val Pro His Ser Leu Pro Phe Lys Val Val
197      275     280     285
200 Val Ile Ser Ala Ile Leu Ala Leu Val Val Leu Thr Ile Ile Ser Leu
201      290     295     300
204 Ile Ile Leu Ile Met Leu Trp Gln Lys Lys Pro Arg
205 305     310     315
208 <210> SEQ ID NO: 3
209 <211> LENGTH: 32
210 <212> TYPE: DNA
211 <213> ORGANISM: Artificial sequence
213 <220> FEATURE:
214 <223> OTHER INFORMATION: antisense primer containing Xba I site
216 <400> SEQUENCE: 3
217 gctctagact ggcccacagc attcagatcc tc
220 <210> SEQ ID NO: 4
221 <211> LENGTH: 28
222 <212> TYPE: DNA
223 <213> ORGANISM: Artificial sequence
225 <220> FEATURE:
226 <223> OTHER INFORMATION: sense primer containing EcoRI
228 <400> SEQUENCE: 4

```

32

RAW SEQUENCE LISTING DATE: 05/10/2001
 PATENT APPLICATION: US/09/403,882A TIME: 11:41:49

Input Set : A:\ES.txt
 Output Set: N:\CRF3\05102001\I403882A.raw

229 ggaattcgcc gaggtcaagc tgcaggag 28
 232 <210> SEQ ID NO: 5
 233 <211> LENGTH: 5
 234 <212> TYPE: PRT
 235 <213> ORGANISM: Artificial sequence
 237 <220> FEATURE:
 238 <223> OTHER INFORMATION: localization sequence targeting the nucleus
 240 <400> SEQUENCE: 5
 242 Lys Lys Lys Arg Lys
 243 1 5
 245 <210> SEQ ID NO: 6
 246 <211> LENGTH: 26
 247 <212> TYPE: PRT
 248 <213> ORGANISM: Artificial sequence
 250 <220> FEATURE:
 251 <223> OTHER INFORMATION: localization sequence targeting the mitochondrion
 253 <400> SEQUENCE: 6
 255 Met Leu Arg Thr Ser Ser Leu Phe Thr Arg Arg Val Gln Pro Ser Leu
 256 1 5 10 15
 258 Phe Arg Asn Ile Leu Arg Leu Gln Ser Thr
 259 20 25
 261 <210> SEQ ID NO: 7
 262 <211> LENGTH: 4
 263 <212> TYPE: PRT
 264 <213> ORGANISM: Artificial sequence
 266 <220> FEATURE:
 267 <223> OTHER INFORMATION: localization sequence targeting endoplasmic reticulum
 269 <400> SEQUENCE: 7
 271 Lys Asp Glu Leu
 272 1
 274 <210> SEQ ID NO: 8
 275 <211> LENGTH: 4
 276 <212> TYPE: PRT
 277 <213> ORGANISM: Artificial sequence
 279 <220> FEATURE:
 280 <223> OTHER INFORMATION: insertion into plasma membrane
 282 <220> FEATURE:
 283 <221> NAME/KEY: VARIANT
 284 <222> LOCATION: (1)..(4)
 285 <223> OTHER INFORMATION: any amino acid
 288 <400> SEQUENCE: 8
 290 Cys Cys Xaa Xaa
 291 1
 293 <210> SEQ ID NO: 9
 294 <211> LENGTH: 6
 295 <212> TYPE: PRT
 296 <213> ORGANISM: Artificial sequence
 298 <220> FEATURE:
 299 <223> OTHER INFORMATION: specific targeting sequences c-terminal

OK
 W-->

RAW SEQUENCE LISTING
PATENT APPLICATION: US/09/403,882A

DATE: 05/10/2001

TIME: 11:41:49

Input Set : A:\ES.txt
Output Set: N:\CRF3\05102001\I403882A.raw

301 <400> SEQUENCE: 9
303 Ser Glu Lys Asp Glu Leu
304 1 5

VERIFICATION SUMMARY

PATENT APPLICATION: US/09/403,882A

DATE: 05/10/2001

TIME: 11:41:50

Input Set : A:\ES.txt

Output Set: N:\CRF3\05102001\I403882A.raw

L:290 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:8

NOTICE TO COMPLY WITH REQUIREMENTS FOR PATENT APPLICATIONS CONTAINING NUCLEOTIDE SEQUENCE AND/OR AMINO ACID SEQUENCE DISCLOSURES

The nucleotide and/or amino acid sequence disclosure contained in this application does not comply with the requirements for such a disclosure as set forth in 37 C.F.R. 1.821 - 1.825 for the following reason(s):

- ☒ 1. This application clearly fails to comply with the requirements of 37 C.F.R. 1.821-1.825. Applicant's attention is directed to these regulations, published at 1114 OG 29, May 15, 1990 and at 55 FR 18230, May 1, 1990.
- ☐ 2. This application does not contain, as a separate part of the disclosure on paper copy, a "Sequence Listing" as required by 37 C.F.R. 1.821(c).
- ☐ 3. A copy of the "Sequence Listing" in computer readable form has not been submitted as required by 37 C.F.R. 1.821(e).
- ☒ 4. A copy of the "Sequence Listing" in computer readable form has been submitted. However, the content of the computer readable form does not comply with the requirements of 37 C.F.R. 1.822 and/or 1.823, as indicated on the attached copy of the marked -up "Raw Sequence Listing."
- ☒ 5. The computer readable form that has been filed with this application has been found to be damaged and/or unreadable as indicated on the attached CRF Diskette Problem Report. A Substitute computer readable form must be submitted as required by 37 C.F.R. 1.825(d).
- ☐ 6. The paper copy of the "Sequence Listing" is not the same as the computer readable form of the "Sequence Listing" as required by 37 C.F.R. 1.821(e).
- ☐ 7. Other: _____

Applicant Must Provide:

- ☒ An initial or substitute computer readable form (CRF) copy of the "Sequence Listing".
- ☒ An initial or substitute paper copy of the "Sequence Listing", as well as an amendment directing its entry into the specification.
- ☒ A statement that the content of the paper and computer readable copies are the same and, where applicable, include no new matter, as required by 37 C.F.R. 1.821(e) or 1.821(f) or 1.821(g) or 1.825(b) or 1.825(d).

For questions regarding compliance to these requirements, please contact:

For Rules Interpretation, call (703) 308-4216

For CRF Submission Help, call (703) 308-4212

PatentIn Software Program Support (SIRA)

Technical Assistance.....703-287-0200

To Purchase PatentIn Software.....703-306-2600

PLEASE RETURN A COPY OF THIS NOTICE WITH YOUR RESPONSE

Raw Sequence Listing Error Summary

ERROR DETECTED SUGGESTED CORRECTION

SERIAL NUMBER: 09/403,882

ATTN: NEW RULES CASES: PLEASE DISREGARD ENGLISH "ALPHA" HEADERS, WHICH WERE INSERTED BY PTO SOFTWARE

- 1 Wrapped Nucleics The number/text at the end of each line "wrapped" down to the next line.
This may occur if your file was retrieved in a word processor after creating it.
Please adjust your right margin to .3, as this will prevent "wrapping".
- 2 Wrapped Aminos The amino acid number/text at the end of each line "wrapped" down to the next line.
This may occur if your file was retrieved in a word processor after creating it.
Please adjust your right margin to .3, as this will prevent "wrapping".
- 3 Incorrect Line Length The rules require that a line not exceed 72 characters in length. This includes spaces.
- 4 Misaligned Amino Acid The numbering under each 5th amino acid is misaligned. This may be caused by the use of tabs
Numbering between the numbering. It is recommended to delete any tabs and use spacing between the numbers.
- 5 Non-ASCII This file was not saved in ASCII (DOS) text, as required by the Sequence Rules.
Please ensure your subsequent submission is saved in ASCII text so that it can be processed.
- 6 Variable Length Sequence(s) contain n's or Xaa's which represented more than one residue.
As per the rules, each n or Xaa can only represent a single residue.
Please present the maximum number of each residue having variable length and
indicate in the (ix) feature section that some may be missing.
- 7 PatentIn ver. 2.0 "bug" A "bug" in PatentIn version 2.0 has caused the <220>-<223> section to be missing from amino acid
sequence(s) . Normally, PatentIn would automatically generate this section from the
previously coded nucleic acid sequence. Please manually copy the relevant <220>-<223> section
to the subsequent amino acid sequence. This applies primarily to the mandatory <220>-<223>
sections for Artificial or Unknown sequences.
- 8 Skipped Sequences Sequence(s) missing. If intentional, please use the following format for each skipped sequence:
(OLD RULES) (2) INFORMATION FOR SEQ ID NO:X:
(i) SEQUENCE CHARACTERISTICS:(Do not insert any headings under "SEQUENCE CHARACTERISTICS")
(xi) SEQUENCE DESCRIPTION:SEQ ID NO:X:
This sequence is intentionally skipped

Please also adjust the "(iii) NUMBER OF SEQUENCES:" response to include the skipped sequence(s).
- 9 Skipped Sequences Sequence(s) missing. If intentional, please use the following format for each skipped sequence.
(NEW RULES) <210> sequence id number
 <400> sequence id number
 000
- 10 Use of n's or Xaa's Use of n's and/or Xaa's have been detected in the Sequence Listing.
(NEW RULES) Use of <220> to <223> is MANDATORY if n's or Xaa's are present.
In <220> to <223> section, please explain location of n or Xaa, and which residue n or Xaa represents.
- 11 Use of <213>Organism Sequence(s) are missing this mandatory field or its response.
(NEW RULES)
- 12 Use of <220>Feature Sequence(s) are missing the <220>Feature and associated headings.
(NEW RULES) Use of <220> to <223> is MANDATORY if <213>ORGANISM is "Artificial" or "Unknown"
Please explain source of genetic material in <220> to <223> section.
(See "Federal Register," 6/01/98, Vol. 63, No. 104, pp. 29631-32) (Sec. 1.823 of new Rules)
- 13 PatentIn ver. 2.0 "bug" Please do not use "Copy to Disk" function of PatentIn version 2.0. This causes a corrupted
"file," resulting in missing mandatory numeric identifiers and responses (as indicated on raw sequence listing).
Instead, please use "File Manager" or any other means to copy file to floppy disk.